

9ª MOSTRA NACIONAL DA
QUALIDADE NO PODER
JUDICIÁRIO

BANCO DE SENTENÇAS

CUIABÁ- AGOSTO DE 2009

Introdução

Identificação

- Do órgão: Tribunal Regional do Trabalho da 23ª Região
- Da unidade: Secretaria de Tecnologia da Informação/Diretoria de Desenvolvimento de Sistemas
- Contato: marcelokobayashi@trt23.jus.br

Nome do trabalho

Banco de Sentenças

Nome do responsável e equipe

- Marcelo Massayuki Kobayashi;
- Wilson dos Santos Batista
- Wandrey Alves Rangão
- João Cláudio Nogueira

Delimitação da ação

Tecnologia da Informação aplicada à atividade judiciária - Criação de um sistema consulta textual de sentenças a ser acessada através do sítio institucional do Tribunal Regional do Trabalho da 23ª Região.

Objetivos

Promover o acesso à base de dados de sentenças produzidas por meio eletrônico através de conectores lógicos e expressões de busca. Como consequência direta, permitir o conhecimento do entendimento dos magistrados que atuam na 23ª Região sobre assuntos e questões relevantes para os advogados bem como dos magistrados, servidores e demais interessados, em outras palavras, a efetiva gestão do conhecimento.

Metas

- Indexação de todas as sentenças produzidas eletronicamente;
- Garantia de atendimento ao disposto nos atos normativos dos conselhos superiores (CSJT e CNJ), especialmente o ato conjunto TST/CSJT 10/2008

Desenvolvimento

Da necessidade

As atividades que envolveram a produção de documentos eletrônicos geralmente aconteceram nos Tribunais como forma de garantir economia de recursos físicos, humanos ou otimizar processos de trabalho. Em geral, o modelo de organização dos documentos se mostra o mesmo independente da mídia (papel ou documento digital). Um documento está associado a uma fase ou ato processual realizado. Nesse contexto a dinâmica de tratamento das informações se manteve a mesma e, se pudéssemos fazer uma analogia histórica, poderíamos afirmar que cada setor compõe um feudo, no que se refere às características de gestão da informação.

A proposta de elaboração de um sistema que permita a consulta às sentenças produzidas eletronicamente surgiu da necessidade de magistrados em consultar, de uma maneira simples e eficiente, decisões sobre determinados assuntos. Trata-se de necessidade importante de colaboração que pode (e deve) ser auxiliada por ferramentas de Tecnologia da Informação. Nesse contexto, torna-se importante o modelo e referência dos sistemas de jurisprudência dos Tribunais, que permitem a pesquisa (por vários critérios associativos) por acórdãos e outros tipos de decisões relevantes para se conhecer o entendimento de determinada corte. Assim formava-se o modelo que serviu de parâmetro em conjunto com a necessidade manifestada.

Das restrições apresentadas

A concepção inicial do sistema já nasceu após o advento do ato conjunto CSJT.TST.GP.SE Nº 10/2008¹ que estabeleceu as diretrizes para a adoção de programas com código aberto no âmbito da Justiça do Trabalho. A referida norma também estabelece que a preferência pela utilização de software dessa natureza deve assegurar também a qualidade e eficiência dos serviços correspondentes. Nesse sentido, o estudo de soluções que pudessem atender aos critérios especificados para uma consulta dessa natureza se deu por critérios de aceitação do mercado, comparativos de ferramentas e soluções disponíveis e com intensas discussões com outras organizações públicas e privadas que estavam desenvolvendo projetos semelhantes. Foram estabelecidos contatos com organizações como o Tribunal Superior do Trabalho (TST), o Tribunal de Justiça do Distrito Federal e Territórios (TJDFT), o Tribunal de Contas da União (TCU) e empresas de consultoria especializada.

Não é exagero afirmar que o aspecto orçamentário foi fator de relevante influência, pois existem atas de registro de preço de soluções de mercado

¹ Disponível em

http://informatica.it.gov.br/pls/portal/PORTAL.www_media.show?p_id=2573255&p_settingssetid=128149&p_settingssiteid=0&p_siteid=134&p_type=basetext&p_textid=2573256. Acessado em 30 de agosto de 2009

como o GOOGLE SEARCH APPLIANCE, já implantado em alguns Tribunais Regionais do Trabalho, como por exemplo, a 4ª e 10ª Regiões. A possibilidade de utilização de outras plataformas como o BR/SEARCH, utilizados nos Tribunais superiores também foi cogitada. Porém os custos envolvidos e o benefício potencial da solução

Além disso, a previsão de crescimento da base de dados de documentos produzidos pelo Tribunal acaba por incluir a solução em uma faixa de licenciamento de porte considerável, uma vez que as ferramentas possuem modalidade de licenciamento que varia em função do número de documentos da base de dados. Se tomarmos como exemplo o GOOGLE SEARCH APPLIANCE, em sua versão mini, o licenciamento possui preço de tabela de U\$ 5.000,00 (cinco mil dólares). Essa versão, embora possua preço competitivo permite a indexação de no máximo 300.000 (trezentos mil) documentos. Para que se tenha uma idéia, este é o volume estimado de documentos produzidos apenas nas unidades judiciárias de 1ª instância da capital em um ano. Ressalto que, na capital, estão sediadas 9 das 26 varas do Trabalho da 23ª Região. Não foram considerados, nesta estimativa, os documentos produzidos pelas unidades administrativas, tampouco na 2ª instância. Além disso, essa versão permite apenas possibilidade de indexação de páginas Internet. Em outras palavras, as sentenças e despachos armazenados em Sistemas Gerenciadores de Banco de Dados ou os documentos que atualmente são armazenados em diretórios de rede não podem ser indexados. O custo para soluções de maior porte acaba por ser bem mais alto. No caso da 23ª Região, existem documentos armazenados em sistemas gerenciadores de banco de dados ORACLE e SYBASE, além de um grande número de arquivos salvos em diretórios de rede.

A indexação de documentos é a operação realizada em sistemas especializados em busca que visa permitir que em uma pesquisa a palavra ou expressão seja encontrada considerando todo o conteúdo do documento (indexação em texto total), ou em parte dele (indexação da ementa de um acórdão, por exemplo).

Porém é importante ressaltar que soluções comerciais citadas de maneira bastante superficial (GOOGLE SEARCH APPLIANCE, por exemplo) acabam por proporcionar vantagens importantes como a eficiência comprovada pelo mercado e a existência de equipamentos (hardware) especializados para essa finalidade, conforme figura abaixo, que acompanham o produto quando da aquisição do produto.



Figura 1 - Exemplo de equipamento de servidor (hardware)

No caso da solução desenvolvida para o banco de sentenças, o custo apresentado é o do equipamento (hardware) utilizado que tem sua utilização compartilhada com outros serviços. O software possui licença de uso livre (software não proprietário).

Funcionamento de um sistema de busca

A busca é um mecanismo de localização de um conjunto de informações na forma de documentos ou páginas através de serviços especializados. O sítio Google é o serviço mais lembrado atualmente. Existem outras formas de acesso à informação, como por exemplo organizações temporais ou categorizadas hierarquicamente. Citamos como exemplo os serviços de compras via Internet como amazon (www.amazon.com), submarino (www.submarino.com.br), buscapé (ww.buscape.com.br), entre outros. É importante lembrar que nesses exemplos citados é possível localizar um item não apenas através da navegação, mas também em buscas categorizadas.

Quando consideramos uma busca por um determinado documento ou um conjunto de documentos, vários critérios e estratégias podem ser utilizados. Uma delas é a categorização, através da criação de espécies de documentos. Uma palavra é relevante se estiver em um documento do tipo sentença judicial, por exemplo. Uma busca por categoria “livros” se torna muito mais eficiente do que buscar uma palavra ou expressão em todos os documentos de um sítio de compras. Assim, a categorização torna a busca mais eficiente desde que haja um modelo de organização informacional adequado.

Outra estratégia que contribui bastante para eficiência em um sistema de busca é a relevância. Tal conceito pode ser definido pela estratégia de ordenação dos resultados de uma consulta disparada em um sistema especializado em busca. A aplicação mais conhecida do conceito é a de apresentar os resultados mais importantes primeiro. O conceito pode parecer simples, mas sua implementação é extremamente complexa, uma vez que um ser humano pode

afirmar que um documento é mais importante do que outro de maneira distinta dependendo do contexto, do momento ou da atividade que deu origem à busca. Para que esse mecanismo funcione de maneira adequada a categorização é fundamental, assim como uma política eficaz de classificação e organização da informação.

De maneira simplificada, uma busca se dá pela seguinte seqüência de operações:

1. Uma expressão de busca (palavras ou frases, podendo incluir ou não operadores lógicos) disparada diretamente em uma página Web ou através de um sistema de informação existente;
2. O processamento dessa expressão de busca, para criação de critérios como operadores lógicos, sinonímia, estratégias de desmembramento de palavras para identificação do radical, podendo incluir resultados com equivalentes em gênero, número ou outros critérios. Essa política de desmembramento de palavras visa identificar a necessidade de, a partir do radical da palavra, localizar também variações como incidências do radical em outro tempo de conjugação verbal, equivalente em feminino, plural, sinônimos, etc. Esta etapa é uma das mais afetadas pela escolha da arquitetura tecnológica da solução principalmente no que se refere à flexibilidade e capacidade de adaptação dessas estratégias aos contextos especializados de utilização;
3. Localização dos documentos através das estratégias definidas no passo anterior;
4. Filtragem e ordenação por relevância;
5. Apresentação para o usuário ou sistema que disparou a consulta.

Trata-se de uma operação simples, mas em larga escala, devemos pensar que, em qualquer Tribunal no país o número de documentos que pode se objeto de buscas (independente da natureza, administrativa, judicial) ultrapassa facilmente a marca de um milhão. Estimativas simples sinalizam que esta marca é atingida em um ano em Tribunais de médio porte. Assim, as buscas devem ser auxiliadas por uma estratégia que permita:

- A pesquisa inicialmente exaustiva (em um grande número de documentos);
- A seleção de um conjunto menor de documentos;
- A busca por um documento específico (dentre os milhões geralmente disponíveis) e;
- Principalmente a possibilidade de encontrar novamente o documento quando necessário.

Tal procedimento pode ser descrito através da figura abaixo:

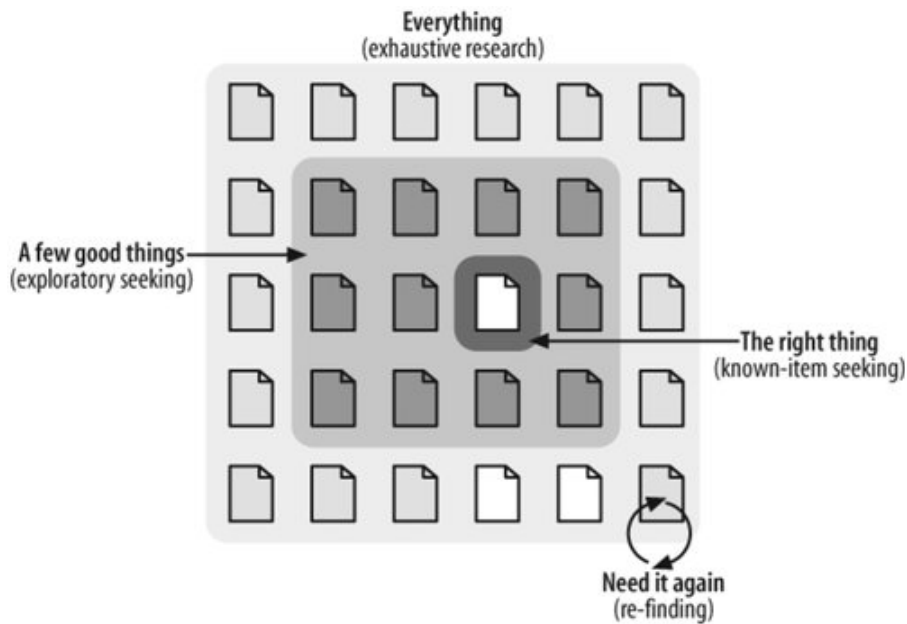


Figura 2 - Mecanismo de busca

Das arquiteturas tecnológicas possíveis

Um dos principais questionamentos apresentados por equipes técnicas, usuários e alta administração tem a ver com a possibilidade de utilização de outras arquiteturas tecnológicas atualmente disponíveis.

Caixa preta

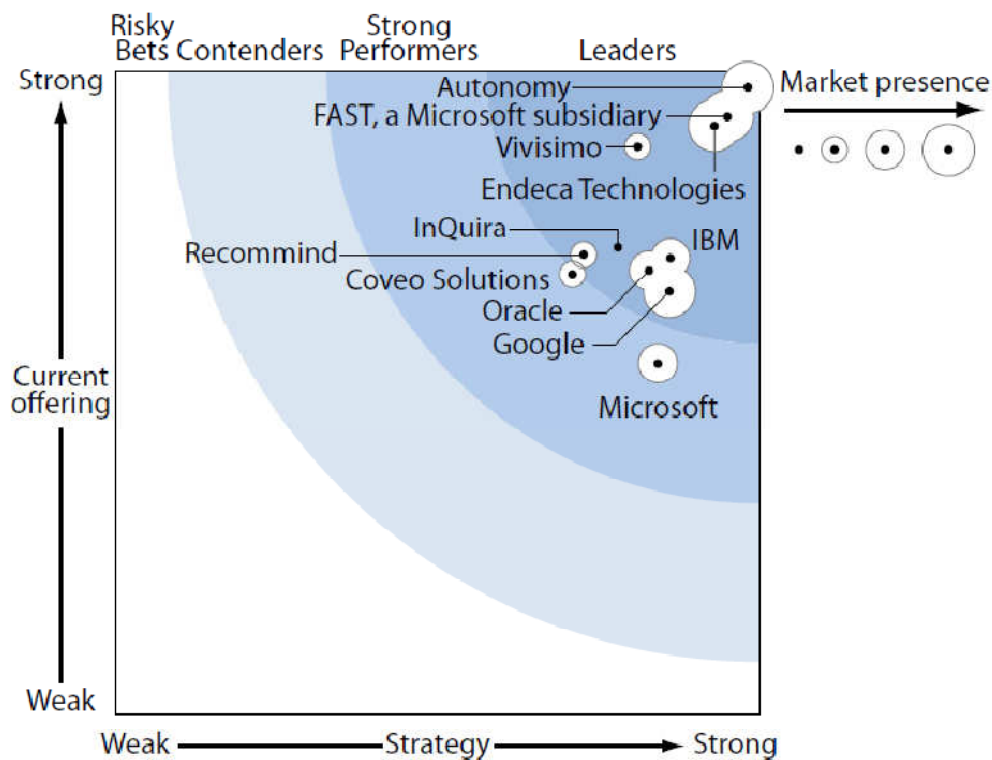


Figura 3 – Produtos líderes do mercado de sistemas de pesquisa (caixa preta)

A primeira delas geralmente é apresentada como uma solução de “caixa-preta”, muito comum no mercado. Essas soluções possuem um alto grau de evolução tecnológica, eficiência nos resultados, garantia de qualidade nas buscas, considerando aspectos como escalabilidade (possibilidade de aumentar capacidade em caso de aumento do número de usuários ou do número de documentos indexados), confiabilidade, alta disponibilidade e performance. Os produtos comerciais com maior presença no mercado geralmente utilizam esta arquitetura.

Essa arquitetura pode ser resumida através do diagrama apresentado a seguir:

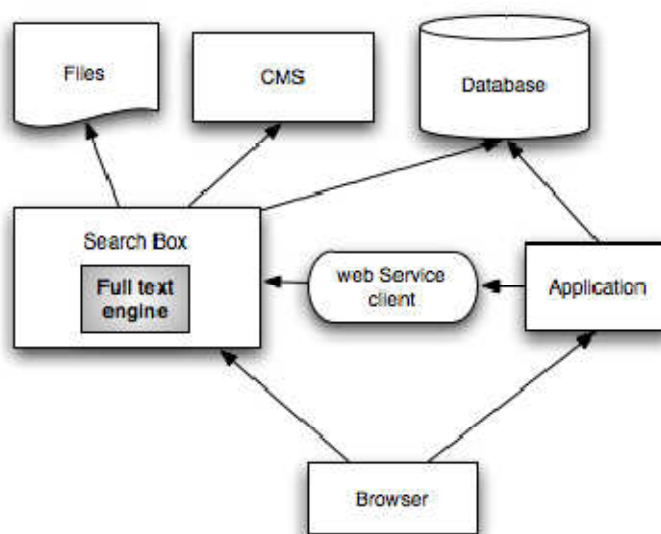


Figura 4 - arquitetura de solução caixa preta

Nesse modelo de solução existe um servidor de consulta (*search Box*), equipamento semelhante ao da figura 2 que recebe as requisições de uma página internet (browser) ou de um sistema (application) através da Internet. Os formatos permitidos são variados e incluem arquivos (files), conteúdo em sistemas de publicação automatizada (cms) e sistemas gerenciadores de banco de dados (database).

As desvantagens dessa solução estão relacionadas à falta de flexibilidade para adaptação a necessidades específicas das organizações e a falta de refinamento na construção de buscas especializadas.

Indexação via SGBD – sistema gerenciador de banco de dados

Outro exemplo de arquitetura possível é a estratégia atualmente utilizada no sistema de jurisprudência da 23ª Região (Tribunal Regional do Trabalho) e também pelo novo sistema de jurisprudência do Tribunal Superior do Trabalho (TST). Trata-se da criação dos índices de consulta (indexação texto total ou parcial) através de recursos dos sistemas gerenciadores de banco de dados, conforme diagrama a seguir:

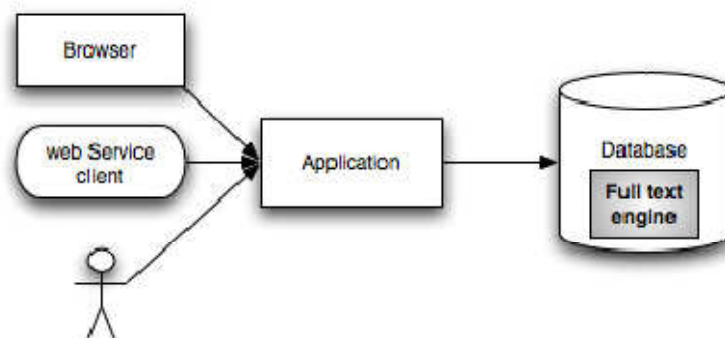


Figura 5 - Consulta através de um SGBD - sistema gerenciador de banco de dados

Neste exemplo os índices ficam armazenados no próprio Banco de Dados (ORACLE nos dois exemplos citados), permitindo uma integração bastante eficiente entre o procedimento de busca e indexação e o processo de armazenamento dos documentos propriamente ditos. O custo também é minimizado pois atualmente as licenças do ORACLE já adquiridas permitem o uso do recurso sem necessidade de novos investimentos. Porém a flexibilidade é mais crítica do que na arquitetura de caixa preta. Apenas documentos armazenados no banco de dados podem ser localizados. Além disso, a dependência força necessariamente a utilização de estratégias e sistemas compatíveis com produtos comerciais que geralmente são incompatíveis com padrões universais e especificações gerais. Acabamos por incidir em uma situação que o ato conjunto TST.CSJT.10/2008 visava evitar: a utilização de software de código aberto visa a utilização de produtos de software “isentos de reservas de direitos proprietários quanto às condições de seu uso, modificação, customização, cessão, liberdade de instalação, plena utilização, acesso ao código fonte, distribuição, padronização”. No caso de produtos comerciais, cada um possui um padrão diferente. Características anteriormente destacadas pela sua importância também não são atendidas por essa solução. Cito a escalabilidade, que se torna bastante prejudicada. Com isso a possibilidade de ampliação da capacidade do sistema para atender a uma demanda maior acaba sendo seriamente prejudicada ou até mesmo impossível.

Bibliotecas de software

A última alternativa arquitetural apresentada neste projeto é baseada no uso de bibliotecas de software. Essas bibliotecas nada mais são do que produtos de software (programas aplicativos) que podem ser integrados em sistemas desenvolvidos pelo Tribunal, por exemplo, ou através de consultas Internet (através do sítio institucional ou serviços informacionais especializados). A grande vantagem dessa abordagem refere-se à flexibilidade. Com o uso dessas soluções é possível, por exemplo, indexar diferentes trechos de um mesmo documento com relevância diferente. Não permitir a consulta sobre partes ou trechos de documentos. Configurar manualmente os índices de maneira a garantir precedência de documentos considerados estrategicamente importantes pela organização. Também é possível integrar a busca com

estruturas de conhecimento a serem criadas como vocabulários controlados² e tesouros³ jurídicos. Por fim é possível agregar inteligência no mecanismo de busca através do estudo da estrutura das palavras (consulta pelo radical, por exemplo), gênero, sinônimos e outras estratégias que tornam o procedimento de busca mais simples para o usuário e mais eficiente por alargar ou restringir o alcance da busca por expressões através dessas estratégias. A arquitetura escolhida para o projeto Banco de Sentenças foi essa. Inicialmente a busca se resume à identificação de critérios lógicos (e, ou) ou por proximidade das palavras na expressão de busca. Brevemente pretendemos agregar estratégias como as citadas acima, também em consonância com o projeto de ARQUITETURA DA INFORMAÇÃO DA JUSTIÇA DO TRABALHO. O modelo da arquitetura é basicamente o demonstrado na figura abaixo:

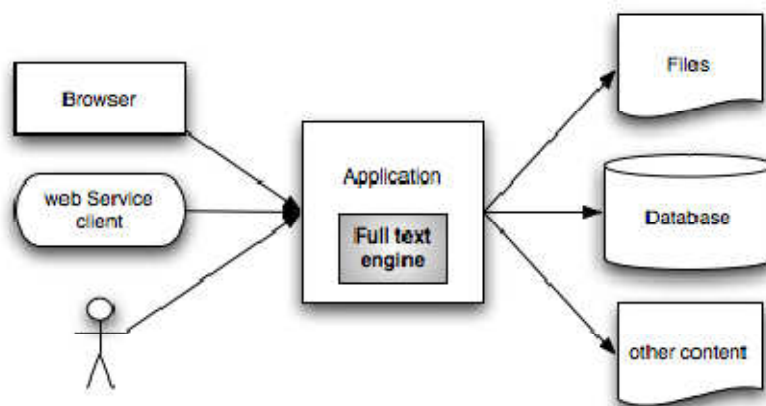


Figura 6 - bibliotecas

Do ponto de vista de aproveitamento dos recursos esta solução é bastante interessante, pois uma vez desenvolvida a biblioteca de software especializada, esta pode ser reaproveitada em todos os sistemas da organização, sem que seja necessário realizar pagar pelo número de documentos existentes ou se prender a uma tecnologia específica de uma empresa. A desvantagem mais citada pela literatura, e reconhecida pelo mercado é a complexidade de criação de uma biblioteca eficiente. Esse fator foi preponderante para que várias empresas procurassem outras alternativas para projetos dessa natureza. No caso do Tribunal Regional da 23ª Região, a estratégia adotada foi a de começar por uma biblioteca simples (desafio já vencido), e agregar funcionalidades à medida que a solução passe a ser utilizada em larga escala, amadureça e as equipes técnicas ganhem mais confiança e experiência na utilização dessa nova tecnologia. A entrada em produção do primeiro serviço

² Um vocabulário controlado pode ser definido como um subconjunto de uma linguagem ou idioma. Existem relações entre palavras e expressões que podem ser de sinonímia ou semelhança (estrito, fino, por exemplo), de especialização (direito → direito público → direito constitucional) ou outras relações que podem ser criadas em diferentes áreas de conhecimento como o direito ou a medicina.

³ Tesouros são espécies de vocabulários controlados. Um tesouro jurídico é um exemplo de vocabulário controlado com associações pertinentes à esfera jurídica

de busca utilizando esse modelo de arquitetura, representada pelo sistema Banco de Sentenças é a demonstração de que tal estratégia se revelou mais apropriada para o contexto do Regional.

Benefícios

Estratégia organizacional

Em uma organização baseada no conhecimento, como um Tribunal, a idéia de compartilhar informações e permitir a colaboração possui bastante relevância no que se refere aos indicadores estratégicos. A tecnologia proporcionada geralmente leva à criação de feudos informacionais ou silos de informação (depósitos de informações departamentais, sem acesso por outros departamentos). Nesse contexto, a utilização de sistemas de buscas como o Banco de sentenças permite o compartilhamento de informações e a efetiva gestão do conhecimento. Não se trata ainda de uma arquitetura informacional ideal, mas um grande passo rumo à colaboração e à criação de ambientes e comunidades práticas.

Tal serviço também demonstra a política de transparência da Administração ao garantir acesso amplo e irrestrito às decisões judiciais produzidas no âmbito de suas unidades jurisdicionais.

Alinhamento

A proposta contempla a utilização de software de código aberto e permite grande flexibilidade para o desenvolvimento de soluções especializadas, conforme interpretação direta do ato conjunto TST.CSJT.10/2008. Outro benefício indireto associado ao alinhamento com as diretrizes emanadas pelo CSJT.

Serviço disponível para magistrados, servidores e público em geral

O projeto, embora demonstre forte influência de novas tecnologias e possibilidades evolutivas, caracteriza-se como um produto disponível aos usuários. A demanda surgiu por conta dos pedidos de vários magistrados que não eram atendidos pelos critérios até então existentes para a localização de documentos de sentenças.

Atualmente o serviço é acessível também pelo público externo, estendendo os benefícios da localização desses documentos ao público em geral, através de link em nosso sítio institucional.

Conclusão

O projeto Banco de Sentenças permitiu que a equipe técnica da STI tivesse contato com tecnologias atuais e em conformidade com o ato conjunto TST.CSJT.10/2008, bem como oportuniza melhorias substanciais em serviços atualmente disponíveis. Como próximas fases do projeto, sugerimos a indexação de outros documentos para que a consulta possa ser realizada também em despachos, acórdãos, documentos administrativos entre outros.

Além disso, a organização do conhecimento através de estratégias como a criação de metadados especializados permite uma política mais eficiente de definição da arquitetura informacional, bem como das políticas de preservação dos documentos.

Ressaltamos que o projeto, embora tenha se materializado em um serviço especializado, já disponível em nosso sítio institucional, também serviu como instrumento para a adoção de tecnologias que podem ser utilizadas em sistemas de informação e em projetos de portais em escala maior, como o portal da Justiça do Trabalho.

Por fim, o conjunto de restrições e fatores que influenciaram os rumos do projeto permitem que haja adaptações a novas realidades e integração com sistemas e soluções desenvolvidas no âmbito deste Regional, da Justiça do Trabalho e outros ramos e órgãos da administração pública. Esperamos que os próximos passos sejam dados no sentido de agregar as funcionalidades descritas como possíveis no item “arquiteturas tecnológicas possíveis – bibliotecas”, quais sejam: a integração da ferramenta com algoritmos inteligentes de processamento das palavras e expressões de busca; e a utilização de um vocabulário controlado especializado. Além disso, a criação de uma arquitetura de informação e definição taxonômica de informações da Justiça do Trabalho, projetos já iniciados no âmbito do CSJT, são iniciativas que podem melhorar substancialmente a utilização da tecnologia utilizada na solução. Assim, com o amadurecimento e a produção de resultados do projeto de arquitetura da informação da JT esperamos evoluir substancialmente o conceito e os serviços de busca descritos neste projeto.